

# Quantum information theory of the Bell-state quantum eraser

Jennifer R. Glick\* and Christoph Adami†

*Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, USA*

(Dated: September 26, 2016)

Quantum systems can display particle- or wave-like properties, depending on the type of measurement that is performed on them. The Bell-state quantum eraser is an experiment that brings the duality to the forefront, as a single measurement can retroactively be made to measure particle-like or wave-like properties (or anything in between). Here we develop a unitary information-theoretic description of this (and several related) quantum measurement situations that sheds light on the trade-off between the quantum and classical features of the measurement. In particular, we show that both the coherence of the quantum state and the classical information obtained from it can be described using quantum-information-theoretic tools only, and that those two measures satisfy an equality on account of the chain rule for entropies. The coherence information and the which-path information have simple interpretations in terms of state preparation and state determination, and suggest ways to account for the relationship between the classical and the quantum world.

## I. INTRODUCTION

Wave-particle duality is an iconic feature of quantum mechanics, one not shared by classical systems in which an object cannot simultaneously have a wave and particle nature. Unraveling the mysteries behind wave-particle duality has occupied the better part of the last century, while significant advances in our understanding have come both from clever experimental approaches as well as theoretical developments. Pivotal experiments were the quantum optical implementations of Wheeler’s [1] delayed-choice experiment [2–4] (which are equivalent in principle to delayed-choice quantum eraser experiments [5, 6]), while the theoretical advances have framed the discussion of the wave-particle duality in terms of a quantum-information-theoretic trade-off between the coherence of the quantum system (or “quanton”) and the information that one may attempt to obtain about the particle path in the interferometer or double-slit experiment (the “which-path” information) [7–11].

The delayed-choice experiments highlight an important feature of this new understanding of wave-particle duality: while as per Bohr’s complementarity principle [12] it is the nature of the experiment that determines whether we shall observe wave- or particle-like behavior, it is clear that the nature of the experiment can be changed after it has already taken place. In other words, the same experiment can retroactively be made to measure wave or particle properties, or anything in between [3, 4, 13]. Such a state of affairs is often greeted with skepticism, as the experiments seem (to some) to imply that the delayed choice of the measurement changes the quantum state retroactively, thus violating causality (see, e.g., the discussion in [13]). The natural interpretation of these results is that a quanton has both particle-like and wave-like properties at the same time, and that

the results of measurements can reveal one or a mixture of those characteristics. Due to the classical nature of the measuring devices, however, they do not—in fact, cannot—reflect the true nature of the quantum state. In the following, we will make these arguments in a strictly information-theoretic setting.

We develop the framework of (possibly delayed) complementary measurements (which-path or which-phase) in terms of a quantum information-theoretic description of the Bell-state quantum eraser, but the formalism is general and applies equally to any situation where measurements are made in parallel on two (and in an obvious extension to several) entangled quantum systems, such as the Garisto-Hardy entanglement eraser [14].

We first describe the ordinary double-slit experiment performed on one half of a Bell state, then the polarization-tagged version where which-path information can be extracted, followed by the quantum erasure procedure. In the next section we describe these steps in terms of classical and quantum information theory that results in an information-theoretic equality that mirrors (and is completely analogous to) the trade-off between distinguishability and visibility of Greenberger and Yasin [7] as well as Englert [8]. The equality involves the coherence of the quanton and the information obtained about its path just as Bagan et al. have recently shown [11], but we do not assume a specific form for the measure of coherence as it emerges naturally from the information-theoretic analysis. The “conservation law” between coherence and information appears simply as a consequence of the chain rule for entropies. We offer conclusions in which we suggest what information is actually encoded in the measurement devices, given that it cannot reflect information about the quantum state.

## II. THE BELL-STATE QUANTUM ERASER

The Bell-state quantum eraser experiment [15], as illustrated in Fig. 1, proceeds as follows. By spontaneous parametric down-conversion (SPDC) [5], a pair of pho-

\* patte399@msu.edu

† adami@msu.edu

photons  $A$  and  $B$  are prepared that are entangled in a Bell state [16]

$$|\Psi\rangle_{AB} = \frac{|h\rangle|v\rangle + |v\rangle|h\rangle}{\sqrt{2}}, \quad (1)$$

where the first and second states refer to photons  $A$  and  $B$ , respectively, and  $|h\rangle, |v\rangle$  are the horizontal and vertical linear polarization states. Photon  $B$ , called the “idler”, is sent in one direction towards a photodetector  $D_B$ . A polarization filter (denoted POL in Fig. 1) can be placed in its path to perform measurements in specific bases. Meanwhile, photon  $A$ , called the “signal”, is sent in another direction towards a double-slit apparatus. Photon  $A$  will pass through the double slit and subsequently be detected by a photodetector  $D_X$  that scans along the  $x$ -axis. From the detector counts of  $D_X$  one can construct an interference pattern as a function of the spatial variable  $x$ . The interference pattern is erased by placing two quarter-wave plates (QWPs) in front of each slit. This tags the path of photon  $A$  and provides path information. See Fig. 1 for a schematic of the experiment.

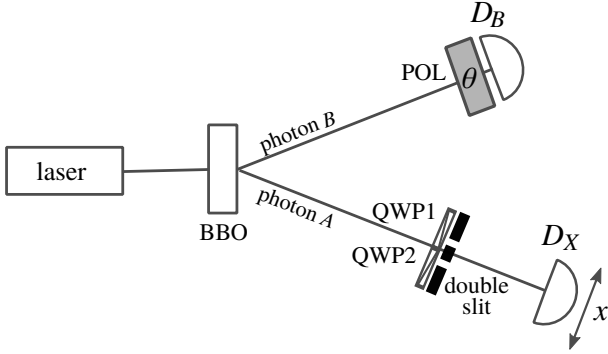


FIG. 1. Schematic of the double-slit Bell-state quantum eraser experiment [15]. After production of the Bell state (1) by type-II spontaneous parametric down-conversion (SPDC) in the BBO ( $\beta$ -barium borate) crystal, photon  $B$  travels along the upper branch to a linear polarizer (POL, set to angle  $\theta$  relative to the  $|h\rangle, |v\rangle$  basis) and detector  $D_B$ . Photon  $A$  travels down to the quarter-wave plates (QWPs) and double slit, and then to detector  $D_X$ , which plays the role of a “screen” by scanning along the direction  $x$ .

### A. Splitting the photon

The full wavefunction describing the entangled pair of photons  $A$  and  $B$  is

$$|\Psi\rangle_{AB} = \frac{|h\rangle_P|v\rangle_B + |v\rangle_P|h\rangle_B}{\sqrt{2}} \otimes |\psi\rangle_Q, \quad (2)$$

where the Hilbert space  $\mathcal{H}_A = \mathcal{H}_P \otimes \mathcal{H}_Q$  of photon  $A$  is composed of polarization  $P$  and spatial  $Q$  degrees of

freedom ( $Q$  is called the quanton). The polarizations of photons  $A$  and  $B$ , entangled in a Bell state, are decoupled from the spatial state  $|\psi\rangle$  of photon  $A$ . We drop the spatial part of photon  $B$  as it remains decoupled throughout. Sending photon  $A$  through the double slit transforms only the spatial states of  $A$  so that Eq. (2) evolves to

$$|\Psi\rangle_{AB} = \frac{|h\rangle_P|v\rangle_B + |v\rangle_P|h\rangle_B}{\sqrt{2}} \otimes \frac{|\psi_1\rangle_Q + |\psi_2\rangle_Q}{\sqrt{2}}. \quad (3)$$

The states  $|\psi_j\rangle$  denote the path of photon  $A$  corresponding to slit  $j$  [17]. The spatial degree of freedom of  $A$ , denoted by  $Q$ , is still independent from its polarization.

Tracing over the states of photon  $B$ , the density matrix describing the spatial modes of photon  $A$  is the pure state

$$\rho_Q = \frac{1}{2} \left( |\psi_1\rangle_Q \langle\psi_1| + |\psi_2\rangle_Q \langle\psi_2| + |\psi_1\rangle_Q \langle\psi_2| + |\psi_2\rangle_Q \langle\psi_1| \right). \quad (4)$$

The expectation value in the position basis  $|x\rangle$  of the screen  $D_X$  yields the probability to observe photon  $A$  at position  $x$

$$\langle x|\rho_Q|x\rangle = p(x) = \frac{1}{2} |\psi_1(x) + \psi_2(x)|^2, \quad (5)$$

where we define the amplitudes  $\psi_j(x) = \langle x|\psi_j\rangle$ . This probability distribution is a coherent superposition and the usual double-slit interference fringes will be observed on the screen. In the appendix we show how the characteristic fringes can be derived from a von Neumann measurement of  $Q$  by the detector  $D_X$ .

### B. Tagging the photon path

To extend this discussion to a quantum eraser experiment, a tagging operation is performed on the two branches of the double-slit apparatus in order to provide information about the path of photon  $A$ . In practice, this is implemented by placing a quarter-wave plate (QWP) in front of each slit. The Jones matrix for a general wave plate (WP) oriented at an angle  $\beta$  (the fast axis) to our coordinate system (in this case,  $|h\rangle$  and  $|v\rangle$ ) is [18, 19]

$$U = \begin{pmatrix} \cos(\frac{\alpha}{2}) + i \sin(\frac{\alpha}{2}) \cos(2\beta) & i \sin(\frac{\alpha}{2}) \sin(2\beta) \\ i \sin(\frac{\alpha}{2}) \sin(2\beta) & \cos(\frac{\alpha}{2}) - i \sin(\frac{\alpha}{2}) \cos(2\beta) \end{pmatrix}, \quad (6)$$

where  $\alpha = \pi/2$  for a QWP. More specifically, the QWP in front of slit 1 (slit 2) has its fast axis at  $\beta = 45$  ( $\beta = -45$ ) degrees, leading to

$$U_{QWP}^{(\pm)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & \pm i \\ \pm i & 1 \end{pmatrix}, \quad (7)$$

where  $U_{QWP}^{(+)} = U_{QWP}^{(1)}$  and  $U_{QWP}^{(-)} = U_{QWP}^{(2)}$  are the matrices associated with slit 1 and 2, respectively. These

transform linearly polarized states  $|h\rangle$  and  $|v\rangle$  into circularly polarized states according to

$$\begin{aligned} U_{QWP}^{(1)} |h\rangle &= \frac{|h\rangle + i|v\rangle}{\sqrt{2}} = |L\rangle, \\ U_{QWP}^{(1)} |v\rangle &= \frac{|v\rangle + i|h\rangle}{\sqrt{2}} = i|R\rangle, \\ U_{QWP}^{(2)} |h\rangle &= \frac{|h\rangle - i|v\rangle}{\sqrt{2}} = |R\rangle, \\ U_{QWP}^{(2)} |v\rangle &= \frac{|v\rangle - i|h\rangle}{\sqrt{2}} = -i|L\rangle, \end{aligned}$$

where  $|R\rangle$  ( $|L\rangle$ ) denotes right-handed (left-handed) circular polarization.

When photon  $A$  passes through the QWPs and the double slit, its polarization becomes entangled with its spatial degree of freedom so that the wavefunction (2) evolves to

$$\begin{aligned} |\tilde{\Psi}\rangle_{AB} &= \frac{1}{\sqrt{2}} \left( \frac{|L\rangle_P |v\rangle_B + i|R\rangle_P |h\rangle_B}{\sqrt{2}} \otimes |\psi_1\rangle_Q \right. \\ &\quad \left. + \frac{|R\rangle_P |v\rangle_B - i|L\rangle_P |h\rangle_B}{\sqrt{2}} \otimes |\psi_2\rangle_Q \right), \end{aligned} \quad (8)$$

where the tilde indicates that the tagging operation has been performed. Grouping together the polarization states of photon  $B$ , we can equivalently express this state as

$$\begin{aligned} |\tilde{\Psi}\rangle_{AB} &= \frac{1}{\sqrt{2}} \left( \frac{|\psi_1\rangle_Q |L\rangle_P + |\psi_2\rangle_Q |R\rangle_P}{\sqrt{2}} \otimes |v\rangle_B \right. \\ &\quad \left. + i \frac{|\psi_1\rangle_Q |R\rangle_P - |\psi_2\rangle_Q |L\rangle_P}{\sqrt{2}} \otimes |h\rangle_B \right). \end{aligned} \quad (9)$$

The entanglement between the two degrees of freedom of photon  $A$  causes the spatial modes  $Q$  to become completely mixed

$$\rho_Q = \frac{1}{2} (|\psi_1\rangle_Q \langle \psi_1| + |\psi_2\rangle_Q \langle \psi_2|), \quad (10)$$

so that interference is no longer observed on the screen. In Fig. (2) we show the entropic Venn diagrams [20] before (a) and after (b) the tagging operation with the QWPs. In these diagrams, the sum of all the entries in a circle add up to the entropy of the subsystem, and the entropy shared between subsystems is indicated in the overlap between circles. Conditional entropies appear in unshared areas of the circle, and can be negative in quantum mechanics [21] (they must be positive if they are classical Shannon entropies). Entropies shared between three systems (the center of the diagrams in Fig. 2) can be negative both in classical and quantum physics [20]. All of the von Neumann entropies  $S(\rho) = -\text{Tr} \rho \log \rho$  can be calculated in a straightforward manner from the density matrix  $\rho_{QPB} = |\tilde{\Psi}\rangle_{AB} \langle \tilde{\Psi}|$  and the marginalized density matrices  $\rho_Q = \text{Tr}_{PB}(\rho_{QPB})$ ,  $\rho_B = \text{Tr}_{QP}(\rho_{QPB})$ , etc.

Before tagging (see Eq. (2)),  $Q$  is completely decoupled from the polarization  $P$  of photon  $A$  and of photon  $B$ , which together are entangled in a Bell state. After tagging (see Eq. (9)), all three variables  $Q$ ,  $P$ , and  $B$  are in a tripartite entangled state. Note that the ternary mutual entropy  $S(Q : P : B)$  vanishes in both diagrams since the underlying density matrix is a pure state [20]. The expression of the ternary shared entropy in terms of sub-system entropies can be read off the Venn diagram in general as  $S(Q : P : B) = S(Q) + S(B) + S(P) - S(QB) - S(QP) - S(PB) + S(QPB)$ , and similarly for any pair-wise shared entropies.

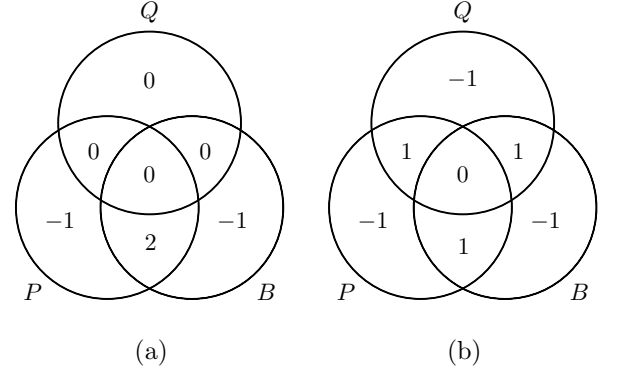


FIG. 2. Entropic Venn diagrams showing the effect of the tagging operation. (a) Before tagging, the polarization states of  $A$  and  $B$  are entangled in a Bell state. (b) After tagging, the spatial degree of freedom of photon  $A$  becomes entangled with the polarization of  $A$  and  $B$  according to Eq. (9).

### C. Erasing the photon path

As is by now well-known [22], it is still possible to extract an interference pattern from the screen data, even when the quanton has been tagged, if we have additional information about the state of photon  $B$ . Suppose we perform a measurement of  $B$  in a basis that is described by an angle  $\theta$  relative to the  $|h\rangle$ ,  $|v\rangle$  basis. For a general change of the basis, the polarization states of  $B$  are written as

$$|v\rangle_B = U_{00}|0\rangle_B + U_{01}|1\rangle_B, \quad (11)$$

$$|h\rangle_B = U_{10}|0\rangle_B + U_{11}|1\rangle_B. \quad (12)$$

For simplicity, we use the following parameterization for the matrix elements  $U_{ij}$  of the rotation operator that transforms  $|h\rangle$ ,  $|v\rangle$  to the new basis spanned by  $|0\rangle$ ,  $|1\rangle$  in terms of the single angle  $\theta$ :

$$U = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (13)$$

Rewriting the states of  $B$  in the basis  $|0\rangle$ ,  $|1\rangle$  and entangling  $B$  with a polarization detector  $D_B$  transforms

Eq. (9) to [20]

$$|\tilde{\Psi}\rangle_{ABD_B} = \frac{1}{2} \sum_{mk} i^m |\psi_m^k\rangle_Q \otimes |m\rangle_P \otimes |kk\rangle_{BD_B}, \quad (14)$$

where  $k = 0, 1$  labels the polarization of  $B$  and  $D_B$  while  $m = 0$  (“ $L$ ”),  $1$  (“ $R$ ”) denotes the polarization of  $A$ , and where we defined the spatial states  $|\psi_m^k\rangle_Q$  of photon  $A$

$$|\psi_L^k\rangle_Q = U_{0k}|\psi_1\rangle_Q - iU_{1k}|\psi_2\rangle_Q, \quad (15)$$

$$|\psi_R^k\rangle_Q = U_{1k}|\psi_1\rangle_Q - iU_{0k}|\psi_2\rangle_Q. \quad (16)$$

These states thus describe the spatial state of photon  $A$  (the quanton) when it has a circular polarization  $m$  and is correlated with photon  $B$  that has polarization  $k$ . Only the states for a given polarization  $m$  are orthonormal

$$\langle\psi_m^{k'}|\psi_m^k\rangle = \delta_{kk'}, \quad (17)$$

$$\langle\psi_L^{k'}|\psi_R^k\rangle = U_{0k'}^* U_{1k} + U_{0k} U_{1k'}^*. \quad (18)$$

Of course, the quanton state  $\rho_Q$  derived from (14) is still completely mixed as in (10) so that no interference can be observed on the screen. However, as long as the erasure angle  $\theta$  is nonzero it is now possible to extract an interference pattern given the outcome of the polarization measurement of photon  $B$  (even if that measurement occurs much later than the measurement of photon  $A$ ).

From the wavefunction (14), we can compute the joint density matrix for photon  $A \equiv QP$  (spatial  $Q$  and polarization  $P$ ) and detector  $D_B$ . Tracing over  $B$  yields

$$\rho_{AD_B} = \frac{1}{2} \sum_k \rho_A^k \otimes |k\rangle_{D_B} \langle k|, \quad (19)$$

where the (orthonormal) states of photon  $A$ , conditional on the state  $k$  of detector  $D_B$ , are  $\rho_A^k = |\phi_k\rangle\langle\phi_k|$  and  $|\phi_k\rangle = \frac{1}{\sqrt{2}} \sum_m i^m |\psi_m^k\rangle_Q \otimes |m\rangle_P$ . The effect of the erasure is contained in the behavior of these states as the measurement angle  $\theta$  is varied. For a measurement in the original basis ( $\theta = 0$ ), detector  $D_B$  prepares photon  $A$  in one of the fully entangled states:  $|\phi_0\rangle \propto |\psi_1\rangle_Q |L\rangle_P + |\psi_2\rangle_Q |R\rangle_P$  and  $|\phi_1\rangle \propto |\psi_1\rangle_Q |R\rangle_P - |\psi_2\rangle_Q |L\rangle_P$ . From these expressions we can infer, with a polarization measurement of photon  $A$ , its path. For instance, outcome  $k = 0$ ,  $m = L$  is associated with the spatial state  $|\psi_1\rangle$  for slit 1. Therefore, polarization measurements of photon  $B$  at  $\theta = 0$  yield full path information and no interference fringes.

On the other hand, for a measurement in the diagonal basis at  $\theta = \pi/4$ , detector  $D_B$  prepares photon  $A$  in one of the completely decoupled states:  $|\phi_0\rangle \propto (|\psi_1\rangle_Q - i|\psi_2\rangle_Q) \otimes (|L\rangle_P + i|R\rangle_P)$  and  $|\phi_1\rangle \propto (|\psi_1\rangle_Q + i|\psi_2\rangle_Q) \otimes (-|L\rangle_P + i|R\rangle_P)$ . Now, a polarization measurement of  $A$  cannot reveal path information, and the coherently summed spatial modes lead to interference fringes. In the appendix, we compute these interference patterns and show their dependence on the erasure angle  $\theta$ . Thus, we see that, regardless of the temporal order of the two polarization measurements, the measurement of  $B$  can

be seen as state *preparation*, while the measurement of  $A$  is state *determination*, that is, extraction of which-path information.

### III. INFORMATION THEORY

#### A. State preparation

We illustrate the quantum erasure mechanism by building on the entropic Venn diagrams in Fig. 2 and constructing the entropic relationships between the random variables  $Q$ ,  $P$ , and  $D_B$  from the joint and marginal entropies associated with Eq. (19). In the basis  $|\phi_k\rangle \otimes |k\rangle$  it is clear that the entropy of (19) is  $S(QPD_B) = 1$ , while the marginal entropies are all  $S(Q) = S(P) = S(D_B) = 1$ . Tracing over detector  $D_B$ , the total entropy of  $A \equiv QP$  is also  $S(QP) = 1$ . The pairwise entropy  $S(PD_B)$  is equal to  $S(QB)$  by the Schmidt decomposition of (14), which in turn is equal to  $S(QD_B)$  by the symmetry between the  $B$  and  $D_B$  states. The remaining pairwise entropy is then computed from the density matrix

$$\rho_{QD_B} = \frac{1}{2} \sum_k \rho_Q^k \otimes |k\rangle_{D_B} \langle k|, \quad (20)$$

where the state of  $Q$  conditional on the outcome  $k$  of detector  $D_B$  is

$$\begin{aligned} \rho_Q^k &= \frac{1}{2} \sum_m |\psi_m^k\rangle_Q \langle\psi_m^k| \\ &= \frac{1}{2} \left( |\psi_1\rangle_Q \langle\psi_1| + |\psi_2\rangle_Q \langle\psi_2| \right. \\ &\quad \left. + i(-)^k \sin 2\theta \left[ |\psi_1\rangle_Q \langle\psi_2| - |\psi_2\rangle_Q \langle\psi_1| \right] \right). \end{aligned} \quad (21)$$

From this expression, we see that for a given outcome  $k$ , the spatial degree of freedom of photon  $A$  is generally no longer mixed (as in Eq. (10)) and has a coherence that is controlled by the sine of the measurement angle. In turn, it is possible to extract interference fringes from the screen. With the Bloch vector  $\vec{a} = -(-)^k \sin(2\theta) \hat{y}$ , Eq. (21) can be expressed as  $\rho_Q^k = \frac{1}{2} (\mathbb{1} - (-)^k \sin(2\theta) \hat{\sigma}_y)$ , where  $\hat{y}$  is a unit vector and  $\hat{\sigma}_y$  is a Pauli matrix. This density matrix varies from a fully mixed state  $|\vec{a}| = 0$  at  $\theta = 0$  to a pure state  $|\vec{a}| = 1$  at  $\theta = \pi/4$ .

To compute the entropy of the block-diagonal matrix in Eq. (20), we first find the entropy of  $\rho_Q^k$ . The eigenvalues of  $\rho_Q^k$  are  $\lambda_{\pm} = \frac{1}{2} (1 \pm |\vec{a}|) = \frac{1}{2} (1 \pm \sin(2\theta))$  and are independent of the index  $k$ , leading to equal entropies  $S(\rho_Q^0) = S(\rho_Q^1)$ . Therefore, the entropy of (20), which is also equal to the entropy  $S(PD_B)$ , is [23]

$$S(QD_B) = 1 + \frac{1}{2} \sum_k S(\rho_Q^k) = 1 + S, \quad (22)$$

where  $S = -\lambda_+ \log \lambda_+ - \lambda_- \log \lambda_-$  is the entropy of (21), and varies from  $S = 1$  at  $\theta = 0$  to  $S = 0$  at  $\theta = \pi/4$ .

The relationship between the variables  $Q$ ,  $P$ , and  $D_B$  is summarized by the entropic Venn diagram in Fig. 3(a). As a result of tagging the path of photon  $A$ , the spatial and polarization modes of  $A$  are entangled, given the state of  $D_B$ , for  $S > 0$  ( $\theta < \pi/4$ ). The amount of entanglement  $S$  varies with the erasure angle  $\theta$  and specifies the degree to which the polarization  $P$  can reveal information about the spatial mode  $Q$ . The non-zero ternary mutual entropy  $S(Q:P:D_B) = 1 - 2S$  indicates that the mutual entropy of  $Q$  and  $P$  can be shared by detector  $D_B$  [24].

### B. State determination

In order to reveal information about the path of photon  $A$ , its polarization is measured after it passes the double-slit apparatus using a detector  $D_A$  in the circular  $|L\rangle, |R\rangle$  basis. After the measurement with  $D_A$ , the wavefunction (14) evolves to

$$|\Psi'\rangle_{AD_ABD_B} = \frac{1}{2} \sum_{mk} i^m |\psi_m^k\rangle_Q \otimes |mm\rangle_{PD_A} \otimes |kk\rangle_{BD_B}. \quad (23)$$

The entropic relations between the variables  $Q$ ,  $D_A$ , and  $D_B$  are computed from their joint density matrix, which is found by tracing out the polarization states of photons  $A$  and  $B$  from (23),

$$\rho_{QD_AD_B} = \frac{1}{4} \sum_{mk} |\psi_m^k\rangle_Q \langle \psi_m^k| \otimes |m\rangle_{D_A} \langle m| \otimes |k\rangle_{D_B} \langle k|. \quad (24)$$

For a set of outcomes  $k$  and  $m$ , the corresponding spatial state of photon  $A$  is  $|\psi_m^k\rangle_Q$ . It is straightforward to show that the entropy of Eq. (24) is  $S(QD_AD_B) = 2$ , while the marginal entropies are  $S(Q) = S(D_A) = S(D_B) = 1$ .

Tracing over the spatial states  $Q$  in Eq. (24), we find that the two polarization detectors are uncorrelated

$$\rho_{D_AD_B} = \frac{1}{2} \mathbb{1}_{D_A} \otimes \frac{1}{2} \mathbb{1}_{D_B}, \quad (25)$$

with a joint entropy  $S(D_AD_B) = 2$  and

$$S(D_A : D_B) = 0. \quad (26)$$

Thus, the measurement with  $D_A$  reveals no information about the state of  $D_B$ . This is not surprising as the QWPs act as a “controlled-NOT” gate on the polarization. Indeed, conditioning on the spatial states of photon  $A$  yields

$$S(D_A : D_B|Q) = S \geq 0. \quad (27)$$

In a sense, therefore, the quanton states encrypt the relationship between the state preparation with  $D_B$  and its readout with  $D_A$ .

Tracing out the states of the polarization detector  $D_A$  in Eq. (24), the joint density matrix  $\rho_{QD_B}$  is unchanged

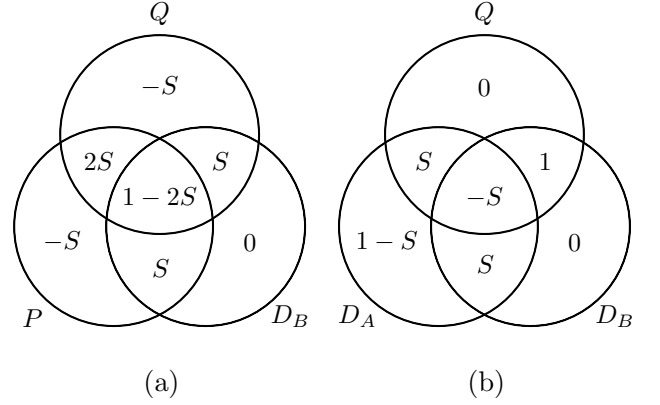


FIG. 3. Entropic Venn diagrams for (a) state preparation with detector  $D_B$  (see Eq. (14)), and (b) state determination with detector  $D_A$  (see Eq. (23)).

from Eq. (20), since the measurement with  $D_A$  does not affect the correlations between  $Q$  and  $D_B$ .

Tracing over the states of  $D_B$  in Eq. (24), the joint density matrix for  $Q$  and  $D_A$  in turn is

$$\rho_{QD_A} = \frac{1}{2} \sum_m \rho_Q^m \otimes |m\rangle_{D_A} \langle m| = \frac{1}{2} \mathbb{1}_Q \otimes \frac{1}{2} \mathbb{1}_{D_A}, \quad (28)$$

where the density matrix of  $Q$ , conditional on the polarization outcome  $m$  of detector  $D_A$ , is

$$\rho_Q^m = \frac{1}{2} \sum_k |\psi_m^k\rangle_Q \langle \psi_m^k| = \frac{1}{2} (|\psi_1\rangle_Q \langle \psi_1| + |\psi_2\rangle_Q \langle \psi_2|), \quad (29)$$

which is independent of the polarization index  $m$ , and is equivalent to the full density matrix  $\rho_Q$ . That is,  $\rho_Q^m = \rho_Q = \frac{1}{2} \mathbb{1}_Q$ , and is a completely mixed state. Finally, the joint entropy of  $Q$  and  $D_A$  is  $S(QD_A) = 2$ . The entropic relationships between the variables  $Q$ ,  $D_A$  and  $D_B$  are summarized by the Venn diagram in Fig. 3(b).

### C. Information-theoretic origins of coherence and path information

From the marginal and joint entropies computed in the previous section, we can construct information-theoretic relationships between the variables  $Q$ ,  $D_A$ , and  $D_B$ . First, as previously noted, detectors  $D_B$  and  $D_A$  share no information

*Coherence.*—The information shared between the preparation with detector  $D_B$  and the quanton  $Q$  (the spatial state of photon  $A$ ) is given by the mutual entropy

$$S(Q : D_B) = 1 - S \leq 1, \quad (30)$$

with  $S$  defined in Eq. (22). We can understand how this entropy depends on the erasure angle  $\theta$  by considering two cases. First, from the joint density matrix  $\rho_{QD_B}$  in

Eq. (20), a measurement of photon  $B$  at an angle  $\theta = 0$  decouples  $Q$  from detector  $D_B$

$$\theta = 0: \quad \rho_{QD_B} = \frac{1}{2} \mathbb{1}_Q \otimes \frac{1}{2} \mathbb{1}_{D_B}, \quad (31)$$

and the conditional state  $\rho_Q^k$  in Eq. (21) becomes a statistical mixture, i.e., interference cannot be observed on the screen. In this case,  $S(Q : D_B) = 0$  and there is no information shared between the two variables. However, increasing the erasure angle to  $\theta = \pi/4$  leads to perfect correlation

$$\theta = \frac{\pi}{4}: \quad \rho_{QD_B} = \frac{1}{2} \left( |f\rangle_Q \langle f| \otimes |0\rangle_{D_B} \langle 0| + |a\rangle_Q \langle a| \otimes |1\rangle_{D_B} \langle 1| \right), \quad (32)$$

where  $|f\rangle_Q = \frac{1}{\sqrt{2}}(|\psi_1\rangle_Q - i|\psi_2\rangle_Q)$  corresponds to a fringe pattern and  $|a\rangle_Q = \frac{1}{\sqrt{2}}(|\psi_1\rangle_Q + i|\psi_2\rangle_Q)$  corresponds to an antifringe (phased-shifted) pattern. Now,  $\rho_Q^0 = |f\rangle_Q \langle f|$  and  $\rho_Q^1 = |a\rangle_Q \langle a|$  are coherent superpositions, i.e., it is possible to extract interference on the screen. At this angle,  $S(Q : D_B) = 1$ . Therefore, the mutual entropy  $S(Q : D_B)$  is related to the coherence of the conditional states  $\rho_Q^k$ , and in turn, to the visibility of interference fringes as we will see below.

*Path information.*—From the joint density matrix  $\rho_{QD_A}$  computed in Eq. (28), the polarization measurement with detector  $D_A$  reveals nothing about the spatial degree of freedom of  $A$  since the joint state is completely decoupled. It follows that the mutual information vanishes

$$S(Q : D_A) = 0. \quad (33)$$

In other words, if we do not know the outcome of the polarization measurement of photon  $B$ , an attempt to measure the polarization of  $A$  after it traverses the double slit and QWPs will not reveal anything about the spatial state of  $A$ . On the other hand, if we do know the state of  $D_B$ , then the conditional mutual information is

$$S(Q : D_A | D_B) = S \geq 0, \quad (34)$$

and varies with the erasure angle  $\theta$ .

To understand the behavior of this quantity as a function of  $\theta$ , consider the state of  $Q$  and  $D_A$ , conditional on the outcome  $k$  of  $D_B$ . According to Eq. (24), this state is  $\rho_{QD_A}^k = \frac{1}{2} \sum_m |\psi_m^k\rangle_Q \langle \psi_m^k| \otimes |m\rangle_{D_A} \langle m|$ . For an outcome  $k = 0$  of a polarization measurement of photon  $B$  at angle  $\theta = 0$ ,

$$\theta = 0: \quad \rho_{QD_A}^0 = \frac{1}{2} \left[ |\psi_1\rangle_Q \langle \psi_1| \otimes |L\rangle_{D_A} \langle L| + |\psi_2\rangle_Q \langle \psi_2| \otimes |R\rangle_{D_A} \langle R| \right]. \quad (35)$$

At this angle, the conditional mutual information is maximal,  $S(Q : D_A | D_B) = 1$ , and it is clear that the state of the polarization detector  $D_A$  is associated with one of the states  $|\psi_j\rangle_Q$ , so that the measurement can reveal information about the path of photon  $A$ . For instance, an outcome  $L$  corresponds to the state  $|\psi_1\rangle_Q$ . As the erasure angle  $\theta$  increases from zero to  $\pi/4$ , the information

we have about the spatial state of  $A$  is reduced to zero, since the density matrix becomes decoupled:

$$\theta = \frac{\pi}{4}: \quad \rho_{QD_A}^0 = |f\rangle_Q \langle f| \otimes \frac{1}{2} \mathbb{1}_{D_A}. \quad (36)$$

At this angle,  $S(Q : D_A | D_B) = 0$ . Therefore, the tagging operation with the QWPs only reveals information about the path of  $A$  as long as we have additional information about the state of photon  $B$  from its polarization measurement with detector  $D_B$ . Thus, Eq. (34) is the correct expression for which-path information. Note further that

$$S(Q : D_A D_B) = 1, \quad (37)$$

which implies that—given the outcomes of both polarization detectors  $D_A$  and  $D_B$ —it is possible to predict *with certainty* the outcome of a direct measurement of the path of photon  $A$ .

#### IV. DISCUSSION

The quantities considered in the previous section can be used to generalize the usual concepts of coherence and path information, allowing us to construct a more fundamental relationship that is derived from information-theoretic principles and a no-collapse model of quantum measurement [20].

Recapitulating the results from the previous section, we know that whether we extract fringes or antifringes from the screen is controlled by the state of  $D_B$ . The visibility of the fringes is, in turn, related to the coherence of the conditional state  $\rho_Q^k$  of  $Q$  in Eq. (21) and the mutual information  $S(Q : D_B)$ . As we have already seen, at angle  $\theta = 0$  ( $\theta = \pi/4$ ) the two variables  $Q$  and  $D_B$  are completely uncorrelated (correlated), the conditional state  $\rho_Q^k$  is a statistical mixture (coherent superposition), and we observe no interference (full interference).

On the other hand, information about the path of photon  $A$  is determined by the correlation between its polarization (via the state of detector  $D_A$ ) and its spatial states. This correlation is computed from the conditional mutual information  $S(Q : D_A | D_B)$ , and must be conditioned on the state of  $D_B$  since  $Q$  and  $D_A$  are otherwise uncorrelated. When  $\theta = 0$  ( $\theta = \pi/4$ ), the variables  $Q$  and  $D_A$  are completely correlated (uncorrelated) given  $D_B$ , so that  $D_A$  can (cannot) reveal path information, and we extract no interference (full interference) from the screen.

These two information-theoretic quantities, namely, the quanton coherence  $S(Q : D_B)$  and the path information  $S(Q : D_A | D_B)$ , are fundamentally linked through the chain rule for entropies

$$S(Q : D_B) + S(Q : D_A | D_B) = S(Q : D_A D_B) = 1, \quad (38)$$

and their sum is conserved throughout the erasure process. This information-theoretic formulation of complementarity generalizes earlier attempts [10, 11, 25] by explicitly referencing the measurement devices. We note

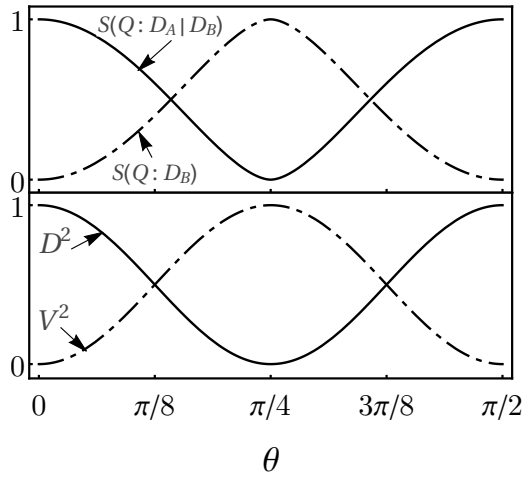


FIG. 4. Top: Relationship between coherence and path information in Eq. (38) for the Bell-state quantum eraser as a function of the erasure angle  $\theta$ . Shown are the information-theoretic quantities for the path information  $S(Q: D_A | D_B) = S$  (solid) and coherence  $S(Q: D_B) = 1 - S$  (dashed). Bottom: The distinguishability  $D$  (solid) and fringe visibility  $V$  (dashed), defined below, as a function of the erasure angle  $\theta$ . When  $\theta = 0$  ( $\theta = \pi/4$ ) there is full (no) path information and no (full) coherence.

that the absence of correlations between detectors  $D_A$  and  $D_B$ ,  $S(D_A : D_B) = 0$ , is crucial to enforce complementarity.

We show in the top of Fig. 4 the coherence and information in Eq. (38) as a function of the erasure angle  $\theta$ . We can compare these expressions to two other measures that are commonly used to discuss the wave-particle duality, namely the distinguishability  $D$  and the visibility  $V$  [7, 8]. In general,  $D^2 + V^2 \leq 1$ , but becomes an equality when the detectors are prepared in pure states. In the particular case we are discussing,  $D^2 = \cos^2(2\theta)$ , while the fringe visibility  $V^2 = \sin^2(2\theta)$ . They are shown in the bottom part of Fig. 4 and exhibit a remarkably similar behavior when compared to the information-theoretic complementarity principle.

From Eq. (38), we can derive additional information-theoretic relations with conditional and mutual entropies. With the definition of conditional mutual information [20],  $S(Q: D_A | D_B) = S(Q | D_B) - S(Q | D_A D_B) = S(Q | D_B)$ , Eq. (38) becomes a relation between a mutual information and a conditional entropy:

$$S(Q: D_B) + S(Q | D_B) = 1. \quad (39)$$

Furthermore,  $S(Q: D_A) - S(Q: D_A | D_B) = S(Q: D_A : D_B) \leq 0$ , so that Eq. (38) becomes

$$0 \leq S(Q: D_A) + S(Q: D_B) \leq 1, \quad (40)$$

where the lower bound comes from the non-negativity of mutual entropies. This can be rewritten in terms of conditional entropies as

$$1 \leq S(Q | D_A) + S(Q | D_B) \leq 2, \quad (41)$$

where we used  $S(Q) = 1$ .

Bagan et al. recently constructed an entropic complementarity relation between coherence and path information in an interferometer [11] using an entropic measure for coherence. For path states with equal probability, orthogonal detector states, and orthogonal measurements, their relation can be written as an equality very similar to ours

$$C_{\text{rel ent}}(\rho) + H(M : D) = 1, \quad (42)$$

where  $C_{\text{rel ent}}(\rho) = 1 - S(\rho)$  is a measure of the coherence of the particle's state  $\rho$  in the interferometer, and  $H(M : D) = S(\rho)$  is the path information, which is the mutual entropy of the path detector states  $D$  and the results  $M$  of probing them with a measurement. The connection to our result (38) is immediately obvious, as  $S(\rho)$  in Eq. (42) is indeed equivalent to our  $S$ , the entropy of the conditional state  $\rho_Q^k$  of photon  $A$  in Eq. (21). However, our measures of coherence and path information emerge naturally from a full information-theoretic analysis and yield more insight into the origins of their complementarity, in particular how the quanton's entropy is distributed among the detectors  $D_A$  and  $D_B$  (as summarized by Fig. 3).

## V. CONCLUSIONS

We prefer to tread lightly when using our results to discuss aspects of quantum theory that have been discussed in a controversial manner since the discussions between Bohr and Einstein concerning these matters [26]. Nevertheless, we believe some statements can be made unequivocally. For example, it is now clear (and has been pointed out repeatedly before us), that a quanton not only carries both particle and wave attributes, but that these quantities are manifested in measurement devices in a fluid manner. In particular, the dynamics of the Bell-state quantum eraser, which allows us to give measurements different “meanings” depending on what state preparation we may choose *after* the state determination has taken place, cannot possibly be consistent with a picture of quantum measurement in which the quantum state is irreversibly projected so as to be consistent with the state of the measurement device. The actuality of not only information erasure, but the production of alternative outcomes via the retroactive manipulation of state preparation, confirms the picture that the quanton wavefunction after measurement continues to carry amplitudes that are *not* consistent with the state of the measurement device.

That the quantum state can be inconsistent with the state of the measurement device should not come as a surprise to practitioners of quantum information science. After all, the idea of the classical measurement, in which the state of the system to be measured is copied onto the state of the measurement device, cannot carry over to quantum mechanics on account of the no-cloning

theorem [27, 28]. Indeed, the central idea of classical measurement—in which the variation of the system is fully correlated with the variation in the device—is impossible for pure quantum states that carry no entropy whatsoever.

Of course, mixed quantum states (pure joint states with a reference state traced out) can carry entropy, and this entropy can be shared with classical measurement devices. This appears to be the case in the construction described here, as the entropy of the quanton is exactly one bit (in the ideal case whose extension is discussed in footnote [16]). If the classical device (here the device  $D_A$ ) cannot carry information about the quanton's state, what information does it reflect? In our view, a classical device's state must be consistent with the state of other classical measurement devices, so as to ensure a causally consistent world. Here, the information  $S(Q : D_A | D_B)$  predicts the outcome of a measurement of the which-path information that would be obtained if a device was placed squarely in the path of the beam. Of course, such a device would record a random outcome (the photon would be found in state  $\psi_1$  half the time), and  $D_A$  would perfectly predict this random outcome. Still, neither of these states predict the quanton's state, which after all is neither here nor there. We are thus forced to admit that our classical devices do not (and *cannot*) reveal to us the quantum reality underlying our classical world [29]. However, experimental (and theoretical) ingenuity has allowed us to be aware of our classical device's deceptions, and shown us the path to perhaps design even more clever schemes to lift the veil from the underlying quantum reality.

## ACKNOWLEDGMENTS

Financial support by a Michigan State University fellowship to JRG is gratefully acknowledged.

## Appendix: Interference patterns

Here we derive the spatial probability distributions for photon  $A$  that is incident on a screen  $D_X$  by modeling the interaction as a von Neumann measurement of the spatial location of photon  $A$ . Expanding the spatial states of  $A$  in terms of the position basis of the screen yields

$$|\psi_j\rangle_Q = \sum_{x=1}^n \psi_j(x) |x\rangle_Q, \quad (\text{A.1})$$

where  $j = 0, 1$  labels each slit. The states  $|x\rangle$  can be discretized into  $n$  distinct locations according to

$$\begin{aligned} |x=1\rangle &= |100\cdots 0\rangle, \\ |x=2\rangle &= |010\cdots 0\rangle, \\ &\vdots \\ |x=n\rangle &= |0\cdots 001\rangle, \end{aligned}$$

which denote the location  $x$  at which a photon is detected by  $D_X$ . Inserting this basis into the expression (14) and performing the measurement of  $Q$  with  $D_X$  (which starts in the initial state  $|x=0\rangle = |0\cdots 0\rangle$ ), we come to

$$|\Psi'\rangle_{AD_X BD_B} = \frac{1}{2} \sum_{xmk} i^m \psi_m^k(x) |xx\rangle_{QD_X} \otimes |m\rangle_P \otimes |kk\rangle_{BD_B}, \quad (\text{A.2})$$

where we define the amplitudes  $\psi_m^k(x) = \langle x | \psi_m^k \rangle_Q$ .

Tracing over  $A$  and  $B$  in Eq. (A.2), we arrive at the (classical) joint density matrix for detector  $D_B$  and the screen

$$\begin{aligned} \rho_{D_X D_B} &= \frac{1}{4} \sum_{xmk} |\psi_m^k(x)|^2 |x\rangle_{D_X} \langle x| \otimes |k\rangle_{D_B} \langle k| \\ &= \frac{1}{2} \sum_k \rho_{D_X}^k \otimes |k\rangle_{D_B} \langle k|, \end{aligned} \quad (\text{A.3})$$

where

$$\rho_{D_X}^k = \sum_x p_k(x) |x\rangle_{D_X} \langle x|, \quad (\text{A.4})$$

are the conditional states of the screen  $D_X$  with corresponding conditional probability distribution

$$p_k(x) = \frac{1}{2} \sum_m |\psi_m^k(x)|^2. \quad (\text{A.5})$$

Tracing out detector  $D_B$  from (A.3) yields the full density matrix for  $D_X$

$$\rho_{D_X} = \frac{1}{2} \sum_k \rho_{D_X}^k = \sum_x p(x) |x\rangle_{D_X} \langle x|, \quad (\text{A.6})$$

where the total probability distribution of the screen is

$$p(x) = \frac{1}{2} \sum_k p_k(x). \quad (\text{A.7})$$

It is straightforward to show that the total probability distribution  $p(x)$  for the screen is completely incoherent due to the cancellation of the cross terms of the two conditional probabilities. That is,

$$p(x) = \frac{1}{2} (|\psi_1(x)|^2 + |\psi_2(x)|^2). \quad (\text{A.8})$$

This distribution describes two intensity peaks on the screen corresponding to the pattern obtained from each slit individually. From the data as a whole (i.e., when we do not know the outcome of detector  $D_B$ ) no interference is observed on the screen.

However, suppose we do know the outcome of the polarization measurement of photon  $B$ . For an outcome  $k$ , the conditional state of the screen  $D_X$  is given by Eq. (A.4). To discern the type of interference pattern the probability distribution (A.5) of this density matrix



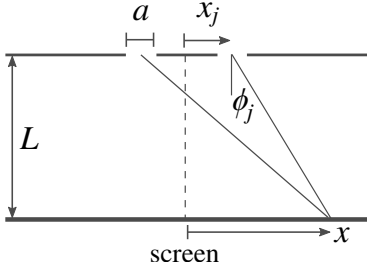


FIG. 5. Geometry of the double slit apparatus. The slit width is  $a$ , the distance from the origin to the center of slit  $j$  is  $x_j$ , the distance from the slits to the screen is  $L$ . The angle from the center of slit  $j$  to point  $x$  on the screen is  $\tan \phi_j = (x - x_j)/L$ .

describes, we rewrite the conditional probability in terms of the original amplitudes  $\psi_j(x) = \langle x | \psi_j \rangle_Q$ , leading to

$$p_k(x) = \frac{1}{2} \left[ |\psi_1(x)|^2 + |\psi_2(x)|^2 + i(-)^k \sin 2\theta \times [\psi_1(x) \psi_2(x)^* - \psi_1(x)^* \psi_2(x)] \right]. \quad (\text{A.9})$$

where we used  $U_{1k} U_{0k}^* + U_{1k}^* U_{0k} = (-)^k \sin 2\theta$ . In general, this expression will describe interference fringes with a visibility that is controlled by the magnitude of the coefficient  $\sin 2\theta$  in front of the cross terms. Let us consider two specific cases of the erasure angle  $\theta$ .

First,  $\theta = 0$  corresponds to a measurement of photon  $B$  in the linear  $|h\rangle, |v\rangle$  basis. In this case, expression (A.9) reduces to an incoherent sum

$$\theta = 0: \quad p_k(x) = \frac{1}{2} (|\psi_1(x)|^2 + |\psi_2(x)|^2), \quad (\text{A.10})$$

and describes two intensity peaks on the screen with no interference. In turn, we have full information about the path of photon  $A$ . Second,  $\theta = \pi/4$  describes a measurement of photon  $B$  in the diagonal  $|\pm\rangle = \frac{1}{\sqrt{2}}(|h\rangle \pm |v\rangle)$  basis. In this case, expression (A.9) becomes a perfectly coherent sum

$$\theta = \frac{\pi}{4}: \quad p_k(x) = \frac{1}{2} |\psi_1(x) - i(-)^k \psi_2(x)|^2. \quad (\text{A.11})$$

That is, the effect of tagging has been erased since the standard double-slit diffraction pattern can be observed. In general, given an outcome  $k$  for detector  $D_B$ , the corresponding state of the screen is  $\rho_{D_X}^k$  with probability distribution  $p_k(x)$ . This leads to fringes with a level of visibility that is determined by the erasure angle  $\theta$ . The distribution for  $k = 0$  is phase shifted relative to  $k = 1$ , so that depending on the state of  $D_B$ , one observes either fringes or antifringes. Therefore, measuring photon  $B$  in a basis characterized by the angle  $\theta$  allows one to tune the visibility of the interference fringes from the standard two-slit diffraction to single-slit diffraction [4].

To explicitly compute the interference patterns, we write the amplitudes of the  $j$ th slit for a photon of wave-

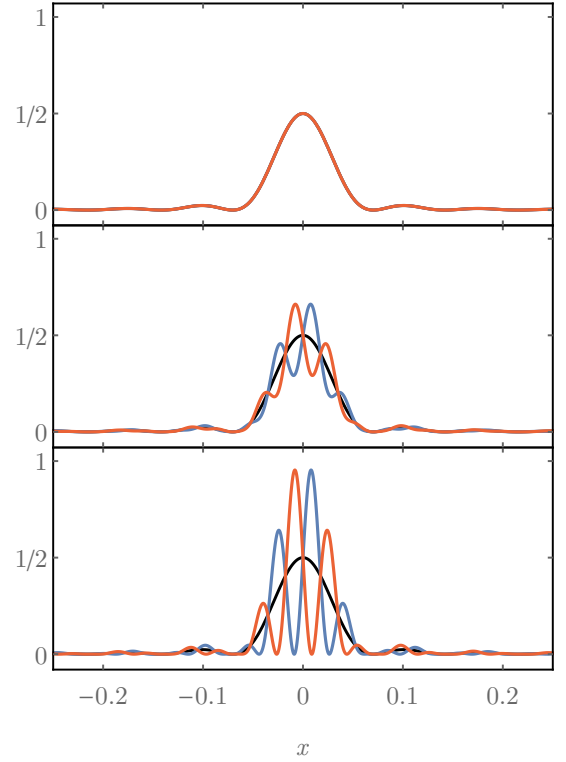


FIG. 6. Conditional probability distributions  $p_k(x)$ , normalized so that the maximum is at 1. The blue (red) oscillations describe the interference pattern  $p_0(x)$  ( $p_1(x)$ ) of photon  $A$  conditional on outcome  $k = 0$  ( $k = 1$ ) of detector  $D_B$ . The black line is the total distribution  $p(x)$  and is the single-slit diffraction result. Parameters for this specific case are  $a = 10 \mu\text{m}$ ,  $d = 20 \mu\text{m}$ ,  $L = 1 \text{ m}$ , and  $\lambda = 702 \text{ nm}$ . Top to bottom: probability distributions for three erasure angles  $\theta = 0$ ,  $\theta = \pi/16$ ,  $\theta = \pi/4$ .

length  $\lambda$  as [30]

$$\psi_j(x) = \sqrt{\frac{a}{2\pi}} \frac{\sin \alpha}{\alpha} e^{-2i\alpha x_j/a}, \quad (\text{A.12})$$

where  $\alpha = \pi a \sin \phi_j / \lambda$ . The geometry of the double-slit apparatus (see Fig. 5) is described by the slit width  $a$ , the distance  $x_j$  to the center of the  $j$ th slit, the angle  $\phi_j = \tan^{-1}((x - x_j)/L)$  from the center of slit  $j$  to the position  $x$  on the screen, and the distance  $L$  from the slits to the screen. For a single slit at the origin, the probability to detect a photon at position  $x$  on the screen is

$$|\psi_j(x)|^2 = \frac{a}{2\pi} \left| \frac{\sin \alpha}{\alpha} \right|^2, \quad (\text{A.13})$$

which is the standard result for single-slit Fraunhofer diffraction. For two slits separated by a distance  $d = x_2 - x_1$ , the amplitudes for each slit are coherently added. In the far-field limit  $L \gg d$ , we can use the approximation  $\phi_j = \phi = \tan^{-1}(x/L)$ . Using the amplitudes  $\psi_j(x)$  in the expression (A.9) for  $p_k(x)$ , leads to the interference patterns shown in Fig. 6. The two patterns  $p_0(x)$  and  $p_1(x)$

are shifted relative to each other on the screen, and the envelope of each pattern is a single-slit diffraction pattern. We show the distributions for three erasure angles:  $\theta = 0, \pi/16, \pi/4$ . For a measurement in the linear  $|h\rangle, |v\rangle$  basis ( $\theta = 0$ ), there is no interference on the screen, since there is full information about the path of photon A. As the erasure angle increases to  $\pi/4$  (a measurement

in the diagonal  $|\pm\rangle$  basis), the oscillations increase to the level of the usual interference pattern for two-slit diffraction. The solid black line is the total distribution  $p(x)$ , which is the full data observed in the experiment, and shows no interference. Only by knowing the outcome  $k$  of detector  $D_B$  can one extract the associated conditional probability distribution  $p_k(x)$  from the full distribution  $p(x)$ .

- 
- [1] J. A. Wheeler, “The “past” and the “delayed-choice” double-slit experiment,” in *Mathematical Foundations of Quantum Theory*, edited by A. R. Marlow (Academic Press, New York, NY, 1978) pp. 9–48.
- [2] V. Jacques, E. Wu, F. Grosshans, F. Treussart, P. Grangier, A. Aspect, and J.-F. Roch, “Experimental realization of Wheeler’s delayed-choice gedanken experiment,” *Science* **315**, 966–8 (2007).
- [3] A. Peruzzo, P. Shadbolt, N. Brunner, S. Popescu, and J. L. O’Brien, “A quantum delayed-choice experiment,” *Science* **338**, 634–7 (2012).
- [4] F. Kaiser, T. Coudreau, P. Milman, D. B. Ostrowsky, and S. Tanzilli, “Entanglement-enabled delayed-choice experiment,” *Science* **338**, 637–640 (2012).
- [5] T. J. Herzog, P. G. Kwiat, H. Weinfurter, and A. Zeilinger, “Complementarity and the quantum eraser,” *Phys. Rev. Lett.* **75**, 3034–3037 (1995).
- [6] Y.H. Kim, R. Yu, S.P. Kulik, Y.H. Shih, and M.O. Scully, “Delayed “choice” quantum eraser,” *Phys Rev Lett* **84**, 1–5 (2000).
- [7] D. M. Greenberger and A. Yasin, “Simultaneous wave and particle knowledge in a neutron interferometer,” *Physics Letters A* **128**, 391–394 (1988).
- [8] B.-G. Englert, “Fringe visibility and which-way information: An inequality,” *Phys. Rev. Lett.* **77**, 2154–2157 (1996).
- [9] P. J. Coles, J. Kaniewski, and S. Wehner, “Equivalence of wave-particle duality to entropic uncertainty,” *Nat Commun* **5**, 5814 (2014).
- [10] R. M. Angelo and A. D. Ribeiro, “Wave-particle duality: an information-based approach,” *Foundations of Physics* **45**, 1407–1420 (2015).
- [11] E. Bagan, J. A. Bergou, S. S. Cottrell, and M. Hillery, “Relations between coherence and path information,” *Phys Rev Lett* **116**, 160406 (2016).
- [12] N. Bohr, “Discussions with Einstein on epistemological problems in atomic physics,” in *Quantum Theory and Measurement*, edited by J. A. Wheeler and W. H. Zurek (Princeton University Press, Princeton, N.J., 1984) pp. 9–49.
- [13] Y. Aharonov and M. S. Zubairy, “Time and the quantum: erasing the past and impacting the future,” *Science* **307**, 875–9 (2005).
- [14] R. Garisto and L. Hardy, “Entanglement of projection and a new class of quantum erasers,” *Physical Review A* **60**, 827–831 (1999).
- [15] S. P. Walborn, M. O. Terra Cunha, S. Pádua, and C. H. Monken, “Double-slit quantum eraser,” *Phys. Rev. A* **65**, 033818 (2002).
- [16] Note that an obvious extension of the present construction is to allow for arbitrary entangled photon pairs, such as for example the q-Bell states [24] in which a parameter  $q$  interpolates between product states ( $q=0$  or 1) and fully entangled states ( $q = 1/2$ ). In that case, certain entries that are 1 bit in the following would be replaced by  $H[q] = -q \log q - (1 - q) \log(1 - q)$ .
- [17] Note that the extension of this framework to allow for  $N$ -path devices is straightforward.
- [18] R. C. Jones, “A New Calculus for the Treatment of Optical Systems, I. Description and Discussion of the Calculus,” *J. Opt. Soc. Am.* **31**, 488–493 (1941).
- [19] E. Collett, *Field Guide to Polarization* (SPIE Press, Bellingham, WA, 2005).
- [20] N. J. Cerf and C. Adami, “Information theory of quantum entanglement and measurement,” *Physica D* **120**, 62–81 (1998).
- [21] N. J. Cerf and C. Adami, “Negative entropy and information in quantum mechanics,” *Physical Review Letters* **79**, 5194–5197 (1997).
- [22] M. O. Scully, B.-G. Englert, and H. Walther, “Quantum optical tests of complementarity,” *Nature* **351**, 111–116 (1991).
- [23] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, 2000).
- [24] C. Adami and N. J. Cerf, “von Neumann capacity of noisy quantum channels,” *Phys. Rev. A* **56**, 3470–3483 (1997).
- [25] W. K. Wootters W. H. Zurek, “Complementarity in the double-slit experiment: Quantum nonseparability and a quantitative statement of Bohr’s principle,” *Physical Review D* **19**, 473–484 (1979).
- [26] J. A. Wheeler and W. H. Zurek, eds., *Quantum Theory and Measurement* (Princeton University Press, Princeton, N.J., 1984).
- [27] W. K. Wootters and W. H. Zurek, “A single quantum cannot be cloned,” *Nature* **299**, 802 (1982).
- [28] D. Dieks, “Communication by EPR devices,” *Physics Letters A* **92**, 271–272 (1982).
- [29] C. Adami and N. J. Cerf, “What information theory can tell us about quantum reality,” *Lecture Notes in Computer Science* **1509**, 258–268 (1999).
- [30] D. Arsenović, M. Božić, and L. Vusković, “Is the classical law of the addition of probabilities violated in quantum interference?” *Journal of Optics B: Quantum and Semi-classical Optics* **4**, S358 (2002).